

UNITED STATES PATENT APPLICATION

for

DYNAMIC BINDING OF NETWORK SERVICES

Inventor:
William Salkewicz

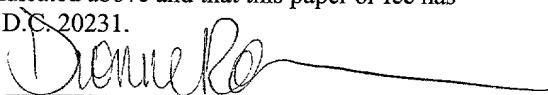
Customer Number 008791
BLAKELY, SOKOLOFF, TAYLOR & ZAFMAN, L.L.P.
12400 Wilshire Boulevard
Seventh Floor
Los Angeles, California 90025-1026
Telephone: (512) 330-0844
Facsimile: (512) 330-0476

Attorney's Docket No. 4906.P001D

"Express Mail" mailing label number: EL 863955655 US Date of Deposit: December 14, 2001

I hereby certify that I am causing this paper or fee to be deposited with the United States Postal Service "Express Mail Post Office to Addressee" service on the date indicated above and that this paper or fee has been addressed to the Commissioner for Patents, Washington, D.C. 20231.

Dionne Robinson
(Printed name of person mailing paper or fee)


(Signature of person mailing paper or fee)

5
Domain Isolation Through Virtual Network Machines And Dynamic Binding of Network Services

Inventor: William M. Salkewicz

BACKGROUND OF THE INVENTION

Field of the Invention

The present invention relates in general to communications networks, and more particularly,
10 to the operation of network devices that can operate in multiple virtual networks simultaneously.

Description of the Related Art

Network Layering and Protocols

15 A communication network provides information resources transfer services that transfer information resources among devices attached to the network. Information resources, as the term is used herein, includes any form of information that can be transmitted over a network for use by or with any end station or network device connected to the network. Information resources, for example, may include computer programs, program files, web pages, data, database information, objects, data structures, program icons, graphics video information or audio information. *Computer Networks and Internets*, Douglas E. Comer, Prentice Hall, 1997, provides extensive information about communication networks.

20 Networks are built from devices or stations called nodes, and the communications channels that interconnect the nodes, called links. A set of nodes and links under one administrative authority is called a network domain. Communication between end stations attached to a network ordinarily is achieved through the use of a set of layered protocols. These protocols are generally described by reference to the Open Systems Interconnection (OSI) computer communications architecture. The standard OSI architecture includes seven layers:
25 application, presentation, session, transport, network, data link and physical. A communication network may employ fewer than the full seven layers. However, the layer 2 and the layer 3 software protocols ordinarily play a prominent role in the transfer of information between interconnected networks and between end stations connected to the networks.

The physical layer is the lowest layer (layer 1) of the OSI model. There are numerous technologies that can be employed to build networks at layer 2. Layer 2 networks can be "connection oriented", meaning that a connection must be established before data can flow between two stations; ATM, Frame Relay, and X.25 are examples of connection oriented 5 layer 2 protocols. Layer 2 networks can also be connection-less, meaning data can be transmitted without establishing any connection in advance; Ethernet and FDDI are two examples of connection-less layer 2 protocols.

In order to provide services useful to end users, the devices in a network must perform higher 10 layer functions to create what are called "virtual networks". The "Internet" is one example of a very popular and public virtual network. The Internet uses the IP protocol to provide the higher layer (layer 3) functions required to support operation of the virtual network. There are many other private (virtual) networks that also uses the IP protocol. The term "internet" with a small "i" is used to differentiate between these less well known private internets, and the 15 very popular and public large "I" Internet. There are many other protocols that can be used to construct virtual networks at layer 3, including IPX, DECnet, AppleTalk, CLNP, etc. There are many other private and public networks using these other layer 3 protocols, either independent of or in conjunction with the IP protocol.

20 Thus, networks can be built at many different layers. Each layer has its own function and its own type of nodes and links. Higher layer networks are built "on top of" lower layer networks. In other words, nodes at a given layer may use the services of the next lower layer to provide links for communication with peer nodes (i.e. nodes at the same layer on other devices). Routers are examples of nodes in a layer 3 network. Bridges are examples of nodes 25 in layer 2 networks.

Network Domains

A network domain as the term is used herein refers to the set of nodes and links that are subject to the same administrative authority. A single administrative authority may administer 30 several networks in separate domains, or several layers of the same network in a single domain, or any combination. There are actually several possible administrative domains in any large virtual network. The boundaries of a network domain can be defined along the lines dividing layers of the protocol stacks. For instance, the same layer 1 physical devices and

physical connections may have several layer 2 network domains layered onto them. These layer 2 domains, in turn, may have one or more layer 3 domains layered on top of them. A network domain may even transcend the boundaries between layers such that a layer 2 network and a layer 3 network may be part of the same network domain.

5

The administration of even a single network domain can be quite complex. Virtual networks have administrative authorities associated with them to control their higher layer functions. The cost of administering a network, physical or virtual, can be enormous, and is often the largest cost item in the operations of a network.

10

When several virtual networks are layered on top of the same layer 2 service or another virtual network, the boundaries between network domains may be somewhat obscure. The boundaries between the domains of the overlaid virtual networks intersect at points where they must share physical or virtual resources. In practice, the administrators of the overlaid virtual networks are very concerned about sharing resources, especially when they are competing commercial entities. Concerns arise about integrity, privacy, and security of data and network control information flowing across the shared resources at the lower layers. The administrators of the underlying networks are called upon to solve complex administrative problems. The costs of administering these networks increases quickly with the number of virtual networks, their size, the complexity and compatibility of their individual policies, and increased demands for security, integrity, and isolation between domains.

Network Devices and Databases

25

The term network device is used here to refer to the collection of mechanisms (e.g. computer and communications hardware and software) used to implement the functions of a station in a network. A network device contains some capacity to store and operate on information in databases in addition to the ability to transmit and receive information to and from other devices on the network. Examples of network devices include but are not limited to routers, bridges, switches, and devices that perform more than one of these functions (e.g. a device that does both routing and bridging).

30

A router is an example of a network device that serves as an intermediate station. An intermediate station is a network device that interconnects networks or subnetworks. A typical router comprises a computer that attaches to two or more networks and that provides communication paths and routing functions so that data can be exchanged between end

stations attached to different networks. A router can route packets between networks that employ different layer 2 protocols, such as Token Ring, Ethernet or FDDI, for example. Routers use layer 3 protocols to route information resources between interconnected networks. Nothing precludes a network device that operates as an intermediate station from 5 also operating as an end station. An IP router for example typically also operates as an end station.

A router can understand layer 3 addressing information, and may implement one or more routing protocols to determine the routes that information should take. A multiprotocol 10 router runs multiple layer 3 protocols such as IP, IPX or AppleTalk for example. A router also be characterized as being multiprotocol if it runs multiple adaptive routing protocols such as RIP, BGP or OSPF all feeding a single IP layer.

The network device router configuration of **Figure 1A** depicts what is often referred to in 15 industry as a multi-protocol bridge/router. In this illustrative example, there are separate databases for three layer 2/3 networking protocols: bridging, IP routing, and IPX routing. The example IP database employs both the OSPF and RIP dynamic routing protocols. Thus, the intermediate station node of **Figure 1A** includes both multiple networking protocols and 20 multiple routing protocols.

A bridge is another example of a network device that serves as an intermediate station. A 25 typical bridge comprises a computer used to interconnect two local area networks (LANs) that have similar layer 2 protocols. It acts as an address filter, picking up packets from one LAN that are intended for a destination on another LAN and passing those packets on. A bridge operates at layer 2 of the OSI architecture.

The term network database will be used to refer to all the control information housed in a 30 network device required to support the device's operation in a set of one or more networks. Each device in a network holds its own network database. In order for the network at large to operate properly, the network databases of all network devices in a network domain should be consistent with each other. The network database control information defines the behavior of its network device. For example, not only might it determine whether the network device will function as a router or a bridge or a switch, but also it will determine the details of how 35 the device will perform those functions.

When a network device is deployed to operate in multiple domains, its network database can become quite complex. The cost of administering the network device increases significantly when the network database is more complex. The cost of administration is already the most 5 significant cost of operating many networks, and the trend toward greater complexity through greater use of virtual networking continues unabated.

The information found in a typical network database includes, but is not limited to, data used to configure, manage, and or monitor operations of:

- 10 Communications Hardware (e.g. layer 1 transceivers/drivers/chips etc.)
- Computer Hardware
- Computer Software
- Layer 2 Addressing
- Layer 2 Connections (Layer 2 interfaces)
- 15 Traffic filter policies
- Bridging (IEEE 802.ID)
- Bridge filters and or policies
- Network (layer 3) Addressing
- Layer 3 Connections (Layer 3 interfaces)
- (Network/layer 3) Address Translation (NAT) policies
- 20 Access Control (e.g. user names and password)
- Access policies (e.g. what user can use what services)
- Routing (IETF RFC 1812)
- Routing Protocols (e.g., BGP, OSPF, RIP, IGRP, etc.)
- 25 Route filters and policies (e.g. route leaking)
- Tunneling
- Tunneling Protocols (e.g., L2TP, GRE, PPTP, etc.)

30 A single network device can operate in one or more (virtual) network domains. For each domain in which a device operates, it needs to store information about that domain in some database form.

Much of the information in a network database must be configured manually; particularly the policy information as it must reflect the administrator's subjective wishes for how the

network should operate. Manual configuration involves human effort, which can become expensive, especially as the number of policies and their complexity increases. Network administrative chores include the assignment of user names, passwords, network addresses or other user identifiers, and configuration of policy databases. This configuration and management may be used to establish traffic filtering policies such as what kind of information payloads will be carried. Traffic and Route filtering policies may be established to determine what paths through the network will be used for each payload carried. Access control policies may be to dictate which users at which end stations have access to which services at other end stations. Security policies may be established to ensure the integrity of the information payloads. Each configured bit of policy somehow finds its way into the network database of the device implementing the policy.

15 *Cisco Router Configuration* by A. Leinwand, B. Pinsky and M. Culpepper, published by MacMillan Technical Publishing, Indianapolis, IN, 1998 provides an extensive treatment of the configuration of the databases of Cisco System routers. This is just one example of a network device database.

Building Virtual Networks

20 The layering of software protocols in accordance with the ISO architecture makes possible the creation of "virtual networks". Virtual networks are to be contrasted with physical networks. Two physical networks which have no physical devices or links in common, can be said to be physically isolated from each other. Physical isolation may be required in order to ensure that a network has the highest levels of security and integrity.

25 Physical networks are defined at layer 1 of the OSI model. Virtual networks, on the other hand, are created at higher layers. It is possible to create multiple virtual networks all sharing common physical resources. A network is definitely virtual if it shares a common physical medium or device, such as an intermediate station, with any other (virtual) network. There are many conventional technologies and many commercially available products which can be used to build many types of virtual networks. For example, virtual circuits are a layer 2 construct that can be employed to create virtual networks.

30 It has been common practice in the industry for phone companies to offer connection oriented layer 1 and 2 services to Internet Service Providers (ISPs), corporations, and residential customers. These customers may build one or more higher layer (layer 3 and above) virtual

networks on top of such publicly available layer 1 and 2 services. The higher layer virtual networks share a common set of layer 1 and 2 services, each having its private set of virtual circuits.

5 A PC or a server are examples of end stations. End stations located at home or business, for example, may connect into an internet through an internet service provider (ISP). There are regional, local and global ISPs. In most cases, local ISPs connect into the regional ISPs which in turn connect into other regional or national ISPs. Figure 1B illustrates an example of connections to an ISP. In the example, home user end stations may connect via modems 10 over dial-up lines to an ISP's router or remote access server (RAS). This data link often runs the PPP (Point-to-Point Protocol) which encapsulates and delivers packets to the ISP's site. Business user end systems may connect to the ISP through leased lines such as T1 lines or T3 lines depending on bandwidth requirements for example. Other examples of typical 15 connection options between home or business users and an ISP include ISDN, T1, fractional T1, various optical media, and xDSL. ISPs may also offer tunnel mode or transport mode services that help businesses set up virtual private networks (VPNs) between remote end stations and virtual dial-up services for remote and mobile end stations.

20 The ISP serves as a conduit for information transmitted between the end stations in the home and other end stations connected to the Internet.

25 A virtual circuit is a dedicated communication channel between two end stations on a packet-switched or cell-relay network. ATM, Frame Relay, and X.25 are all different types of virtual circuit based networking technologies. A virtual circuit follows a path that is programmed through the intermediate stations in the network.

30 There are permanent and switched virtual circuits. A permanent virtual circuit (PVC) is permanent in the sense that it survives computer reboots and power cycles. A PVC is established in advance, often with a predefined and guaranteed bandwidth. A switched virtual circuit (SVC) is "switched" in the sense that it can be created on demand analogous to a telephone call. Both PVCs and SVCs are "virtual" circuits in that they typically are not allocated their own physical links (e.g. wires), but share them with other virtual circuits running across the same physical links.

“Tunneling” is one mechanism for building higher layer networks on top of an underlying virtual network. Tunneling has already gained acceptance in the industry and several technologies are either in operation or under development. Some of the tunneling protocols used in IP networks for example include L2TP, GRE, PPTP, and L2F. There are many other 5 Tunneling technologies used in IP and other protocols.

Referring to **Figures 2A-2B**, there are shown network graphs representing two illustrative networks. Network A is represented by three nodes (NA1, NA2, and NA3), and three links (LA1, LA2, and LA3). Network B is represented by four nodes (NB1, NB2, NB3, and NB4) 10 and four links (LB1, LB2, LB3, and LB4). As used herein, the term node may represent any end station or intermediate station, and the term link means any connection between nodes. If these are physical nodes and links, Networks A and B are physically isolated from each other. If these are virtual (circuit) links which actually depend on a shared physical medium, then the two (virtual) networks are said to be virtually isolated from each other.

15 Illustrative Networks A and B each may be part of different network domains. Independent administrative control may be exercised over each of the Network A and B domains, for example, through the configuration and management of intermediate stations such as bridges and routers.

20 Referring to **Figures 2A** and **2B**, it will be appreciated that the independent administration of the Network A and Network B domains may result in incompatible policies as between the two domains. This is not a problem provided that the domains remain isolated from each other. Referring to **Figure 3**, however, there is shown a network graph of Network C which 25 comprises Networks A and B joined by link LJ. The isolation between Networks A and B, whether physical or virtual, is lost when they are joined in Network C. This joining of the two Networks A and B may create challenges to the administration of combined Network C. For example, despite the joining of the two networks, there still may be a need to apply different or even conflicting policies to each of Networks A and B. In essence, the 30 administrative challenge is to maintain the administrative integrity of the Network A domain and the administrative integrity of the Network B domain despite the fact that both of these networks are part of Network C and are no longer physically isolated from each other.

Figure 4. is an illustrative drawing of a segment of a single physical medium capable of carrying multiple information flows, each in its own virtual circuit (or channel). The physical medium may for instance be a cable or a wire or an optical fiber. The segment shown is carrying four independent information flows on four different virtual circuits; VC1, VC2, VC3, and VC4. These virtual circuits, for example, may be implemented using X.25, ATM, Frame Relay, or some other virtual circuit (or channelized) service.

Figure 5 is an illustrative drawing representing an example of two virtual networks (VN1, and VN2) each made up of two independent network segments (VN1.1 and VN1.2 for VN1, and VN2.1 and VN2.2 for VN2). All segments connect to shared physical network resources. In this example, the shared network resources of Figure 5 provide a virtual circuit service. A virtual circuit connection to an end station or intermediate station connection to a virtual circuit is called a virtual channel connection (VCC). VN1 connects at VCC1 and VCC4; and VN2 connects at VCC2 and VCC3. The shared network resources also provide virtual circuit service that connect VCC1 and VCC4 so as to join VN1.1 and VN1.2 into VN1 and so as to join VN2.1 and VN2.2 into VN2.

Figure 6 is an illustrative drawing that provides additional details of some of the physical constituents of the virtual networks of Figure 5. An intermediate station labeled VN1.1.VCC1 in VN1 connects segment VN1.1 to the VC service at VCC1. An intermediate station labeled VN1.2.VCC4 in VN1 connects segment VN2 to the VC service at VCC4. The VC service connects VCC1 to VCC4, linking VN1.1 to VN1.2 at the virtual circuit level. More specifically, physical media segments PM2, PM1 and PM5 and intermediate stations IS-A and IS-B provide the requisite physical infrastructure upon which the virtual circuit connection linking VN1.1 and VN1.2 is carried. This first virtual circuit connection serves as a network link between the VN1.1.VCC1 and VN1.2.VCC4 intermediate stations, to create one virtual network from the two segments VN1.1 and VN1.2.

Similarly, VCC2 and VCC3 are connected by the virtual circuit service, which connects intermediate stations VN2.1.VCC2 and VN2.2.VCC3, joining the VN2.1 and VN2.2 segments to form the virtual network labeled VN2. More particularly, physical media segments PM4, PM1 and PM3 and intermediate stations IS-A and IS-B provide the virtual connection linking VN2.1 and VN2.2. The second virtual circuit connection serves as a

network link between the VN2.1.VCC2 and VN2.2.VCC3 intermediate stations, to create one virtual network from the two segments VN2.1 and VN2.2.

5 **Figure 7** is an illustrative drawing shows the logical or higher level view of the two virtual networks VN1 and VN2 of **Figures 5 and 6**. It will be appreciated from the view of **Figure 6** that they share physical resources, and it will be appreciated from the view of **Figure 7** that they are logically or virtually separate.

10 In the illustrative example of **Figure 8**, two virtual networks are layered on top of a third virtual network. The sharing of a common set of physical or virtual network resources by several virtual networks increases the challenges of maintaining isolation and security of the individual virtual networks. Nevertheless, end user requirements for information resources, technology advances, economics, politics, and regulations surrounding the networking industry are driving commercial, private and government entities to share common physical and virtual network infrastructure. Therefore, there are ever increasing demands imposed upon network administrators, and vendors of networking equipment.

15 20 25 30

In the illustrative drawing of **Figure 8**, three separate network domains intersect at node IN1: i) that of the Internet itself (including or subsuming that of the underlying VC service supporting the Internet); ii) that of private virtual network VN1; and iii) that of private virtual network VN2. This intersection of three network domains creates the potential for the kinds of administration and policy challenges discussed above. It will be noted that these networks are represented by different network “clouds” that symbolize the multifarious nodes and links in each of the networks.

25

30

The illustrative drawing of **Figure 8** illustrates an example of building two virtual networks on top of another virtual network similar to the previous example in **Figures 5, 6 and 7**. As before, the virtual networks being overlaid are each composed of two segments. Using a tunneling protocol or some other higher layer (layer 3 or above) mechanism, connections are made between nodes IN1.1 and IN1.2 to form a link to tie the two segments of VN1 together. This link is shown as T1 in **Figure 9** and **10**. Link T2 is similar, formed between nodes IN2.1 and IN2.2, to tie the two segments of VN2 together. The logical view of the two virtual networks in **Figure 9** is shown in **Figure 10**, which bears a very strong resemblance to **Figure 7**. The important difference to note between the examples is that in **Figure 7** a layer 2

VC network was used as the underlying network shared resources, and in **Figure 10** another virtual network was used as the underlying network shared resources; specifically, a tunneled service across the Internet. Thus, it will be appreciated that different virtual networks can be formed in different layers using the same underlying physical (or virtual) network resources.

5

Connections are established between nodes at the edge of the segments where they interface or connect to the shared (Internet) resources which are analogous to the virtual circuits in **Figures 5, 6, and 7**. These may be tunneled connections, or connections built using some other (connection-less) technology.

10

If we assume T1 and T2 are tunnels, the network databases of IN1.1, IN1.2, IN2.1, and IN2.2 would be augmented with data structures to manage the tunneling protocol at those endpoints, and the links made up from the tunnels. The network database of IN1.1 of **Figure 8** is depicted in **Figure 11** which highlights the “Tunneling Database” and the “IP Database”.

15
20
25

Network Database Organization

If we examine the information in the network database for IN1, we will see that it should include configuration and policy information for three separate domains. Furthermore, since the information from the three domains must all coexist in the same physical device, there should be some way to structure the information and control its usage, such that the IN1 device operates correctly in all three domains. If all information for the device IN1 were stored in one monolithic form as is done conventionally, in addition to all the policies for each domain, inter-domain policies would also be required to ensure that information should be kept private to its own domain.

25

The illustrative drawing of **Figure 12** is a generalized drawing of a conventional monolithic structure for a database that can be used to implement node IN1 of **Figure 7**. The drawing depicts, in a conceptual fashion, an example of the typical organization of information within such a device. The illustrative device includes a first interface attached to VN1.1, a second interface attached to VN2.2 and a third interface attached to the Internet as the shared network resources. To illustrate the complexities in the database design, assume that both the virtual networks being overlaid on the Internet are also (private) IP networks (internets).

30

Therefore all three networks/domains operate using the IP protocol, each having its own independent IP information to be stored in IN1's network database.

The database includes information such as rules used to articulate and implement

5 administrative policies. The policies as articulated in the information and rules, for example, may include security rules, restrictions on access and dynamic routing protocols. In this illustrative router, the policy information and policy rules used to control the layer 3 IP protocol routing for all three networks are included in a single monolithic database.

10 However, as explained above, different network domains may have different or perhaps even conflicting policies. In order to provide at least some degree of isolation, additional and complicated "inter-domain" policy mechanisms must be added to manage the conflicts between policies on similar data from different domains. These mechanisms are configured and managed by an administrative authority. The dotted lines in **Figure 12** represent the points at which these inter-domain policy mechanisms would be introduced. The policies would attempt to divide the monolithic network database of node IN1 into three separate domain-specific sections. These dotted lines indicate that separation policy mechanisms are implemented, to provide at least some isolation of the information pertaining to VN1 from the information pertaining to VN2, and also from the information pertaining to the Internet (i.e. shared network resources).

15 It will be appreciated that the complexity and difficulty in defining and administering the policy mechanisms used to achieve isolation can be great. There is potential for a wide range of policies to be defined between domains. Everything in the spectrum from almost complete

20 openness and sharing of all information between domains, to the other extreme of not sharing anything at all are possible. Certain pieces of a domain's database may want to be kept private (e.g. access control policy configuration), while other parts are shared to some extent (e.g. summarized routing and addressing information). The types of data, and the extent to which they can all be shared, are all subject to restriction through definition of inter-domain

25 policies.

30

If we consider each boundary between a pair of domains (i.e. each dotted line through the network database of IN1 in **Figure 12**) as a separate policy object, it will also be appreciated that the number of policy objects increases much faster than the number of domains. If D is

the number of domains, then P, the number of policy objects can be calculated approximately as:

$$P = (D(D - 1))/2$$

Thus, the number of policy objects increases approximately as (a proportion of) the square of the number of domains. In other words, the number of policy objects ordinarily increases much faster than the number of domains, especially as the number of domains gets large.

Another challenge in the administration of virtual networks arises because home or business end station users may wish to change the nature of their connections to the network from time to time. For instance, an end user may wish to utilize a more expensive higher bandwidth connection for business use and a less expensive lower bandwidth connection for home or personal use. Alternatively, for instance, an end user may wish opt to receive a video transmission on a higher bandwidth connection while still receiving other transmissions on lower bandwidth connections. An end user may even wish to change the ISP that he or she uses. Unfortunately, these changes often require intervention by a network administration authority to change the higher level binding between the end user station and the network. More specifically, the binding (or association) between the layer 2/1 virtual circuit service and a layer 3 intermediate device is 'hard', not dynamic, and the higher layer interface generally must be reconfigured by a network administrator to change the binding.

Thus, there has been a need for improved organization of network domain databases and improvements in the ability of a network user to change network domain. The present invention meets these needs.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1A is a generalized diagram of a multi-protocol bridge/router;

Figure 1B is an illustrative example of the topology of and connections to an internet

25 service provider (ISP);

Figures 2A and 2B are network graphs of two illustrative example networks;

Figure 3 is a network graph of an illustrative network in which the networks of

Figures 2A and 2B are joined;

5 **Figure 4.** is an illustrative drawing of a segment of a single physical medium capable of carrying multiple information flows, each in its own virtual circuit (or channel);

10 **Figure 5** is an illustrative drawing of two virtual networks each made up of two independent network segments;

15 **Figure 6** is an illustrative drawing that provides additional details of some of the physical constituents of the virtual networks of **Figure 5**;

Figure 7 is an illustrative drawing which shows the logical or higher level view of the two virtual networks VN1 and VN2 of **Figures 5 and 6**;

20 **Figure 8** is an illustrative drawing that shows that the Internet can provide the shared network resources of **Figures 5 and 6**;

15 **Figure 9** is an illustrative drawing that shows tunneling through the Internet to provide the shared resources of **Figures 5 and 6**;

Figure 10 is a logical or high level view of the two virtual networks of **Figure 9**;

25 **Figure 11** is a generalized illustrative drawing of the organization of node IN1 to achieve tunneling;

Figure 12 is a conceptual drawing of one possible router configuration that can be used to implement intermediate node IN1 of **Figure 7**;

30 **Figure 13**, is a generalized block diagram of a network device that instantiates multiple virtual network machine routers in electronic memory in accordance with one embodiment of the invention;

Figure 14 is a generalized block diagram of a network device that instantiates a virtual network machine with multiple layer 2 sub-interface data structures and multiple layer 3 interfaces and binding data structures that associate layer 2 sub-interface data structures and layer 3 interfaces;

35 **Figure 15** is a generalized block diagram of the network device of **Figure 14**, except that one binding data structure has been removed and another binding data structure has been created;

500002000100
Figure 16 is a generalized block diagram of a network device that implements a virtual network machine router and a virtual network machine bridge;

5 **Figure 17** is a generalized block diagram of the network device as in **Figure 16**, except that one binding data structure has been removed and another binding data structure has been created;

10 **Figure 18** is a generalized block diagram of the network device of **Figure 14**, except that one binding data structure has been eliminated and another binding data structure has been created;

15 **Figure 19** is a generalized block diagram of an network device which comprises a computer which instantiates multiple virtual machines in accordance with an embodiment of the invention;

20 **Figure 20** is a generalized block diagram of the network device of **Figure 19** except that one binding data structure has been removed and another binding data structure has been created;

25 **Figure 21** is a generalized block diagram of a subscriber management system in accordance with a presently preferred embodiment of the invention; and

30 **Figure 22** is a generalized block diagram of one of the virtual routers and associated data structures of **Figure 21** in accordance with a present embodiment of the invention.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

35 The present invention comprises a novel apparatus and method for managing operation of network devices that can operate in multiple virtual network domains. The following description is presented to enable any person skilled in the art to make and use the invention, and is provided in the context of particular applications and their requirements. Various modifications to the preferred embodiments will be readily apparent to those skilled in the art, and the generic principles defined herein may be applied to other embodiments and applications without departing from the spirit and scope of the invention. Thus, the present invention is not intended to be limited to the embodiments shown, but is to be accorded the widest scope consistent with the principles and features disclosed herein.

Virtual Network Machines

A Virtual Network Machine (VNM) as the term is used herein to describe the collection of processes and mechanisms that operate on a network device to implement the functions of a node in a virtual network. The preferred embodiment for the VNM is as a set of computer programs and related data structures encoded in electronic memory of a network device and used to operate on information, consuming some portion of a network device's computer and memory storage capacity. The functionality of a virtual network machine can be that of a router, bridge or switch, depending on what is configured in its network database. The native resources of a network device include its processor(s), memory, I/O, communication hardware and system software. The native resources of a network device, for example, may include peripheral devices or even a server computer which may, for instance, provide information about end user privileges or virtual network configurations.

Referring to the illustrative drawing of **Figure 13**, there is shown a generalized block diagram of a new structure for the network database of node IN1 from **Figures 8 and 12** in accordance with one embodiment of the invention that supports creation of multiple virtual network machines. In this case, the networks device IN1 supports three virtual network machines VNM0, VNM1 and VNM2. In the embodiment of **Figure 13**, assuming again that all three virtual networks operate using the IP protocol, each virtual machine implements the functionality of an IP router, each operating in its own network domain. Each virtual network machine is allocated a portion of the device's native resources. Each virtual network machine runs the IP protocol stack. Each virtual network machine stores its address, policy and control information separately from the others. Thus, each virtual network machine can operate independently of the other virtual network machines, even though it shares native computer resources with the other virtual network machines. This virtual network machine based organization of information therefore provides greater isolation between network domains.

Each virtual machine has its own network database that contains its control information. VNM0 has a network database that causes it to operate as a router that routes information within the Internet network domain. VNM1 has a network database that causes it to operate as a router that routes resource information within network domain VN1. VNM2 has a network database that causes it to operate as a router that routes resource information within network domain VN2.1. *High Speed Networks, TCP/IP and ATM Design Principles*, by

William Stallings, Prentice Hall, 1998 provides detailed discussion of router functions and the functions of other network devices.

The VNMs of **Figure 13** may employ multiple different kinds of layer 1 (physical) media to attach to one or more networks. In a presently preferred embodiment, these physical connections include ATM OC-3c/STM1, ATM DS-3/E3, DS-3 Clear Channel, HSSI and 10/100 Base-2 T TX. Resource information is transmitted across these physical connections such as phone lines, DSL or ADSL for example to and from VNM0, VNM1 and VNM2 using layer 2 (data link) protocols. There are layer 2 LAN (local area network) technology and layer 2 WAN (wide area network) technology protocols. Examples of LAN technologies include Ethernet and IEEE 802.3, Fast Ethernet, Token Ring and Fiber Distributed Data Interface. Examples of WAN technologies include Asynchronous Transfer Mode (ATM), Frame Relay, X.25, Point-to-Point (PPP), Integrated Services Digital Network (ISDN) and High-Level Data Link Control (HDLC). Intermediate stations communicate with each other using layer 3 protocols. Layer 3 protocols include Internet Protocol (IP), AppleTalk and Inter Packet Exchange (IPX). Thus, for example, VNM0, VNM1 and VNM2 each employ one or more layer 3 protocols to communicate with other stations of the network(s) to which they are attached.

Thus, the three virtual machines and the different network domains associated with them are isolated from each other in the network device intermediate station of **Figure 13**, and the task of exercising administrative control can be simplified significantly. Since there is no monolithic database that must be maintained to control information transfers across all of the networks to which the three VNMs are attached, the task of administering each database is simplified.

The virtual network machine based organization also simplifies the administration, lowering the cost of operating all three networks. The organization of information along network domain boundaries eliminates the notion of information from two domains residing under a single monolithic structure, and thereby eliminates the need to define inter-domain policies to manage the separation of information within a monolithic database structure. The separation policy mechanisms represented by the dotted lines cutting through the database of **Figure 12** are gone, and a whole set of administrative chores disappears with them. There will be no need to define the complicated inter-domain policies, and no cost associated with administering them. The amount of information that needs to be configured by the

administrators is greatly reduced in size and complexity using this method of database organization.

Other benefits can be realized through greater efficiencies in the implementation of such network devices that are possible with this method of network database organization. Further 5 efficiencies are realized through the elimination of the complicated inter-domain policies in virtually all functions of the device. Essentially, each of the virtual machines VNM0, VNM1 and VNM2 operates a separate/independent network device, performing networking functions its own domain.

Dynamic Binding

10 The drawing of **Figure 14** shows another illustrative embodiment of the invention. The IP network device of **Figure 14** implements a router that includes three network interfaces NIF3-0, NIF3-1 and NIF3-2. The network device also has a layer 1/2 connection to an Ethernet service. The network device also has a layer 1/2 connection to a virtual circuit service. An Ethernet service sub-interface data structure Eth1 provides the layer 2 Ethernet connection such as sub-interface data structure provides the layer 2 VCC1 connection. For example, the VCC1 sub-interface data structure of **Figure 14** may be kept in a table that identifies all virtual circuit connections, each defining the encapsulation protocol, the packet or cell, data compression technique and the particular layer 2 protocol used on that circuit. The Ethernet sub-interface data structure may include the Ethernet address of the local connection and other parameters to control transmit and receipt of information on the Ethernet segment. A binding data structure B3-0 binds the Ethernet sub-interface data structure to NIF3-0. A binding data structure B3-2 binds the VCC1 sub-interface data structure to NIF3-2. The Ethernet and VCC1 sub-interface data structures are labeled with the prefix "sub" because they are layer 2 constructs which are below the layer 3 interface constructs in the ISO scheme. 20 25

30 Referring to **Figure 14**, binding data structure B3-0 establishes a layer 2/3 connection between the Ethernet sub-interface data structure and NIF3-0, and binding data structure B3-2 establishes a layer 2/3 connection between VCC1 sub-interface data structure and IF3-2. Binding data structure B3-0 causes information transferred across the Ethernet connection to be processed through to NIF3-0. An IP Forwarding/Routing database controls routing of the

information out the correct interface. Binding data structure B3-2 causes the information transferred across the VCC1 connection to be processed through NIF3-2.

The VCC1 sub-interface data structure instantiates a virtual circuit connection to the network device of **Figure 14**. A virtual circuit connection such as that in **Figure 14** can be created in 5 accord with any of several technologies. A sub-interface data structure like that in **Figure 14** stores the network device's identity of the virtual circuit attached to it. Many virtual circuits can be established across a single physical connection, and many virtual circuits can be connected to a single network device.

Figure 15 depicts the same intermediate station as in **Figure 14**, except the binding B3-0 has 10 been eliminated, and binding B3-1 has been created. Binding B3-1 associates the Ethernet sub-interface data structure Eth-1 with interface NIF3-1. Interface NIF3-2 remains bound to the sub-interface data structure VCC1. The interface NIF3-0 is not bound to any layer 2 construct. It should be noted that an unbound interface construct generally would represent a mis-configuration in a typical earlier intermediate station.

Figure 16 depicts yet another illustrative embodiment of the invention. The network device of **Figure 16** implements an IP router function and a bridging function. The router includes 15 two interfaces NIF4-1 and NIF4-2. The bridge includes a bridge interface BR4-0. A network database that implements the bridge function includes a list of network stations reachable through each of the bridge's interfaces. The network device also has a layer 1/2 connection to an Ethernet service. The network device also has a layer 1/2 connection to a virtual circuit service VCC1. An Ethernet service sub-interface data structure Eth1 provides information 20 concerning the Ethernet connection such as a VCC1 sub-interface data structure provides information concerning the VCC1 connection. A binding data structure B4-0 binds the Ethernet sub-interface data structure to NIF4-0. A binding data structure B4-2 binds the VCC1 sub-interface data structure to NIF4-2. NIF4-1 is unbound.

Figure 17 depicts the same network device as in **Figure 16**, except the binding B4-0 has 25 been eliminated, and binding B4-1 has been created. Binding B4-1 associates the Ethernet sub-interface data structure with interface NIF4-1 of virtual router VM4. Interface NIF4-2 remains bound to the sub-interface data structure VCC1. The interface BR4-0 is not bound to any layer 2 construct. These changes in binding effectively redefines the service available on the Ethernet segment from a bridged or layer 2 service, to a routed or layer 3 service. In a

presently preferred embodiment of the invention, these bindings can be changed without reconfiguration of any other interface construct or circuit construct. In a typical earlier intermediate station, the bindings between the higher and lower layers are implicit, and a change in the implicit bindings applied to the bridge and router interface constructs typically 5 would have required a modification of these interface constructs. A present embodiment of the invention does not require such modification.

Figure 18 depicts the same network device as in Figure 14, except the binding B3-0 has been eliminated and binding B3-2A has been created. Binding B3-2A associates the Ethernet 10 sub-interface data structure with the NIF3-2 interface. Binding B4-2 associates the VCC1 sub-interface data structure with NIF3-2. Interfaces NIF3-0 and NIF3-1 are unbound. This change in bindings causes both the Ethernet and the virtual circuit lower layer services to be associated with a single higher layer IP construct, NIF3-2.

Figure 19 shows a network device which comprises a computer which instantiates multiple 15 virtual network machines VNM5 and VNM6. VNM5 implements IP router functionality. It includes network interfaces NIF5-0 and NIF5-1. VNM6 also implements IP router functionality. It includes two interfaces NIF6-0 and NIF6-1. The network device of Figure 20 19 has two layer 1/2 connections to a virtual circuit service. Sub-interface data structure VCC1 instantiates one of the connections to the device. Sub-interface VCC2 instantiates the other connection to the device. A binding data structure B5-0 binds the VCC1 sub-interface data structure to NIF5-0 of VNM5. A binding data structure B6-2 binds the VCC2 sub-interface data structure to interface NIF6-1 of VNM6. VNM5 and VNM6 each use the IP protocol suite to communicate with other stations of the network(s) to which they are attached.

Figure 20 depicts the same network device as in Figure 19, except the binding B5-0 has 25 been eliminated and binding B6-0 has been created. The binding B6-0 data structure associates VCC1 sub-interface data structure with NIF6-0 of VNM6. Binding data structure B6-1 binds sub-interface data structure VCC2 to NIF6-1. Neither of the VNM5 interfaces NIF5-0 and NIF5-1 are bound.

In Figures 14 to 20, bindings are shown as data structures connected to other data structures 30 by line segments. In one preferred embodiment, the line segments each represent a pair of bi-directional pointers; the first pointer points from the binding to the higher or lower layer data

structures and the second is opposite the first, pointing from the higher or lower layer data structure to the binding data structure. Alternatively, the binding could be implemented as indices or identifiers in a table, for example. Dynamic binding is accomplished by creating and or deleting binding data structures and or changing the values of the pointers or indices 5 so they operate on different data structures. It will be appreciated that actual changing of the bindings can be accomplished through entries in a command line interface to the network device or automataically by snooping the information flow through the device, for example.

The illustrative drawing of **Figure 21** is a generalized block diagram of a subscriber management system in accordance with a presently preferred embodiment of the invention. 10 A subscriber is a user of network services. The system includes a computer with layer 1/2 connections to subscriber end stations and with layer 1/2 connections to network devices that provide access to other networks.

The system can form a multiplicity of layer 1/2 subscriber end station connections. In a present embodiment, the layer 1/2 connections to subscriber end stations include virtual circuit connections. The system memory stores a multiplicity of sub-interface data structures that instantiate the multiplicity of virtual circuit connections through which subscriber end stations communicate with the subscriber management system.

The system instantiates in memory a plurality of virtual network machines. Each VNM of the embodiment of **Figure 21** implements the functionality of a router. There are nine 20 illustrative VNM routers shown in **Figure 21** labeled VNM₁-VNM₉. Each VNM router includes interfaces in its database. Each VNM router runs at least one layer 3 protocol suite. Each VNM router may run one or more adaptive routing algorithms. The interfaces of each VNM router provide access to a network that is isolated from the networks accessed through the interfaces of the other VNM routers. For example, the interface to VNM₄ provides 25 layer 3 access to the network that includes ISP#2. The interface to VNM₅ provides layer 3 access to the network that includes Corporate-Private-Network#A. The interface to VNM₆ provides layer 3 access to the network that includes ISP#4. The networks with ISP#2, Corporate-Private-Network#A and ISP#4 are isolated from each other. The databases 30 associated with VNM₄, VNM₅ and VNM₆ to control access to networks across these respective interfaces. Each of these three VNM databases can be administered separately.

In operation a subscriber might establish a point-to-point connection with the subscriber management system. A server that runs software that runs authentication, authorization and accounting protocols (AAA) searches for a record that identifies the user. Authentication is the process of identifying and verifying a user. For instance, a user might be identified by a 5 combination of a username and a password or through a unique key. Authorization determines what a user can do after being authenticated, such as gaining access to certain end stations information resources. Accounting is recording user activity. In the present embodiment, AAA involves client software that runs on the subscriber management system and related access control software that runs either locally or on a remote server station 10 attached to the network. The present embodiment employs Remote Authentication Dial-In User Service (RADIUS) to communicate with a remote server. An example of an alternative AAA protocol is Terminal Access Controller Access Control System (TACACS+). RADIUS and TACACS+ are protocols that provide communication between the AAA client on a router and access control server software.

15 The subscriber record includes information concerning the network to which the subscriber's virtual circuit connection should be bound. Typically, the subscriber will employ a PVC. Based upon the information in the subscriber record, a binding data structure, like that described in reference to **Figures 14 to 20**, will be created to associate the sub-interface data structure that instantiates the PVC in the subscriber management system memory with the 20 interface to the VNM router that accesses the network identified for the subscriber in the subscriber record.

Moreover, the subscriber record may provide multiple possible binding options for the subscriber. For instance, the subscriber may specify the creation of a binding that is which is to be employed during business hours and which binds the subscriber to VNMr-5 which 25 provides layer 3 network access to the Corporation-Private-Network#. The same record may specify another binding which is to be employed only during non-business hours and which binds to VNM#4 which provides layer 3 network access to ISP#2. thus, the bindings can be changed. They are dynamic.

30 Various modifications to the preferred embodiments can be made without departing from the spirit and scope of the invention. Thus, the foregoing description is not intended to limit the invention which is described in the appended claims in which: